MODIS
SCIENCE DATA SUPPORT TEAM
PRESENTATION

January 10, 1992

AGENDA

1. Action Items
2. TEG Status
3. Guidelines for MODIS Team Members Science Algorithms
4. MODIS SDST Facilities: Functional Requirements and Near-Term Plans (under separate cover)

ACTION ITEMS:

08/30/91 [Lloyd Carpenter and Team]: Draft a schedule of work for the next 12 months. Include primary events and milestones, documents to be produced, software development, MAS support, etc. STATUS: Open. Due date 09/27/91.

12/06/91 [Liam Gumley]: Investigate a cataloguing scheme for the MAS data. Consider the Master Catalogue, PLDS and PCDS. STATUS: Open. Due date 02/14/92.

12/06/91 [Liam Gumley, Tom Goff, Ed Masuoka]: Develop a plan for storing and distributing MAS data. STATUS: Open. Due date 02/14/92.

01/03/92 [Ed Masuoka]: Check on the UCAR "copyright" as a first step in standardizing an SDST software copyright statement for code sharing. Check with legal. (Being persued) STATUS: Open. Due date 02/14/92.

01/03/92 [Team]: Get a "long arc" of high-resolution AVHRR data for study of time correlations of pointing errors. STATUS: Open. Due date 02/14/92.

01/03/92 [Team]: Check on the set of software engineering tools available in Code 530 to see if any of these would be of use to the SDST. STATUS: Open. Due date 02/14/92.

Tom Goff's Status
for
9 January, 1992

TGoff on GSFC mail,
teg@LTPIRIS2.GSFC.NASA.GOV,
or (301) 982-3704

* **Computer Source Code Availability** - The file dumping utility that has been mentioned over the last several meetings (FDUMP) is been used as a sample case for the determination of legal problems when posting source code (or executables) for public access. We wish to retain a small amount of control over the code while allowing its free distribution, but not be responsible for subsequent actions. For example, we do not want to be responsible for a crashed operating system whether or not a user modifies the code for his/her own purposes. The investigation of copyright protection, code ownership, warranty, "as is" clauses, etc is being pursued with Goddard legal counsel.

* **Anonymous FTP facilities** - The anonymous ftp account on the ltpiris2 computer is in the process of being updated into a group ownership to allow MODIS group users to post files to this area.

* **Science Code Portability** - I have been designated to attempt a port of Mike King's cloud top computer code to a UNIX machine to determine some of the problems that are likely to be encountered. They source code and data sets have been successfully ported to the LTP cdc910b16 iris computer. The program compiles with no errors but need lots of work to execute. Here are some items that need to be considered:

    - This code is rather poorly documented: no references to science documents for the techniques involved, no type casting, almost all variables are three (3) characters in length, nothing is indented, mixed case was not used, comments contain FORTRAN equations instead of descriptions, cryptic English, etc.

    - The code reads the logical unit numbers via the IBM JCL. This requires a different, totally incompatible, technique on non-IBM MVS machines. IBM JCL batch techniques are not used in the rest of the computer industry. Data files are normally designated by file name within the program, not a logical to physical translation at run time.

    - Several GSFC IBM-only functions are used: fread, fwrite, rewind, remtms, and posn. These need to be written or otherwise obtained on the target machine for this program to be successfully linked. Writing subroutines that perform these functions would not be difficult except for the error return address resolution.

    - A library of the above mentioned subroutine functions should be generated to allow porting of routines among the various machine to be encountered in the MODIS time frame.

* The file dumping utility (FDUMP) and the companion character replacement utility (REPLACE) will be placed on the MODIS anonymous ftp site as soon as some legalities are straightened out!

DISCUSSION ITEMS

* A management decision to distribute software contrary to the GSFC COSMIC management directive.

DRAFT
(Incomplete)

MODIS Science Data Support Team (SDST)

Guidelines for MODIS Team Members Science Algorithms

Adapted from "Principles of Numerical Modeling with Examples from Atmospheric Radiation" by Warren J. Wiscombe

DRAFT
(Incomplete)

Table of Contents

DRAFT
(Incomplete)

1 Introduction

The objective of these guidelines is to facilitate the porting, integration, testing, documentation, and maintenance of MODIS Science Team Members algorithms for the generation of MODIS data products on an operational basis. The MODIS Science Data Support Team (SDST) provides support for the development of the MODIS science data processing system. The SDST will generate a plan for the development, validation, integration, operational testing, documentation, maintenance, modification, and configuration management of the MODIS science data processing algorithms. These guidelines are designed to assist the MODIS Science Team Members in preparing their algorithms for this process.

The primary goals for MODIS algorithms developed under these guidelines are:

anticipation and prevention of error
simplicity
easy changeability
portability.

Conditions which will defeat these goals are:

careless organization
bad or non-existent documentation
badly chosen algorithms
precision problems
ill-conditioning.

2 Goal #1: Prevention of Error

Take advantage of existing good debugged modules. Never write a module to do something that someone else has already done.

Test for invalid and implausible input variables in all modules, no matter how unlikely it is that the module will be used incorrectly.

Test non-trivial modules independently of the algorithm itself:

for reasonableness of results
against limiting cases with analytic solutions
against published results where possible
exercising every logical branch
for possible failure.

Document every input, output, and important internal variable, including its units.

Document code blocks with references to specific published equations wherever possible.

Avoid devious or unnecessarily complicated algorithms.

Every time you find and "fix" a code problem, put in an error trap to catch it if it occurs again.

Avoid "quick fixes" of serious errors.

Don't make frequent small changes to the algorithm.

Use plotted output routinely to spot subtle errors.

Walk through the code line by line, being sure each variable has been defined.

Force variables to be "strongly typed".

Always turn on "array bounds checking" during development.

Take advantage of good, debugged existing code before developing new code.

3 Goal #2: Simplicity

Develop the algorithm code and data structures from the top down. The main module should do little more than establish constants and data structures and call other modules.

Choosing simple, elegant data structures leads to dramatically simpler code.

Develop a logical flow that imitates natural human thought processes.

4 Goal #3: Easy Changeability

"Localize" the algorithm so that the range of influence of any change is small.

Keep variables localized by defining them close to the place where they are used.

Keep the logic localized by avoiding long jumps (GO TO statements).

Use short modules that communicate only through argument lists. Avoid global variables, COMMON blocks and EQUIVALENCE.

Strive for maximum generality so the program can be changed with minimum typing.

Dimension all arrays with symbolic constants.

Use logical or integer flags in the input to select among various options.

Concentrate system-dependent plotting, file manipulation, etc. in a few well-chosen general modules that call low-level primitives to do the work.

Develop a comprehensive set of test cases, and develop a driver program which runs the algorithm and compares its results with tabulated answers.

Document the algorithm with descriptive material at the beginning of each module, not just comments within the body.

5 Goal #4: Portability

Portability can be considered as an aspect of "easy changeability", but in the case of algorithms for the MODIS products, portability is of primary concern.

The best guideline is to adhere strictly to standard FORTRAN or C, and resist the temptation of system-dependent features.

Quoting Kernighan and Plauger:

> "We deal with portability by specifying a small set of primitive operations for accessing the environment. All of our programs are written in terms of these primitives, so operating system dependencies are confined to a handful of procedures and functions. Programs that use them can then move to any system where the primitives can be implemented."

6 Documentation

Comments are interspersed throughout a module. Documentation, on the other hand, occurs only at the beginning of a module. It is specific, and it should be revised to reflect changes. Completeness, not brevity, is the main consideration. The documentation should, at a minimum, consist of:

purpose of the module

definition and units of the input and output variables

description and units of the important internal variables

method used, if applicable

references to the literature

notes and warnings (conscious design limitations).

Thorough and complete input variable definitions and descriptions are especially important. Units, type, dimensions, special cases, upper and lower limits, usage examples, default values, and interrelations with other input variables should all be provided.

7 Readability

The most important target of a program is a human; therefore programs should be readable. The single best criterion of program quality is readability.

8 Structure

It is much easier to debug a well-structured program. Unrestrained branching with GO TO statements creates unreadable, unstructured spaghetti-like programs.

Although structuring may slow down initial coding somewhat, the major effort of debugging and maintenance are dramatically speeded up by good structure.

9 Modularization

The aim should be to design a program consisting of loosely coupled modules, each of which does a simple task. Avoid combining several functions together arbitrarily. Modules should be limited to 1 or 2 pages where possible.

10 Software Tools

Software tools are modules of modest size that:

solve a general problem, not a special case;

are nearly perfect, having gone through a considerable shakedown process before being released; and

are user-friendly enough that programmers will prefer them to building their own.

The idea is to create more complex programs mainly by combining tools in different ways. Tools can be divided into categories as utility, numerical analysis, or scientific.

11 Test Drivers

No algorithm should be distributed without comprehensive test drivers that explore all the major branches of the algorithm.

Test drivers should be developed with the same structured, top-down approach used with the algorithms.

Test drivers should "failure-test" the model by pushing it into regimes where trouble is expected.

"Correct answers" should be included in the driver code.

A minimalist approach should be taken. There is no need to generate large print-outs.

Each problem in the driver should be completely independent of the others.

Test drivers should be critically evaluated in the same way as the algorithms.

MODIS SDST FACILITIES:

FUNCTIONAL REQUIREMENTS AND NEAR-TERM PLANS

MODIS SDST Facilities:

Functional Requirements and Near-Term Plans

1.0 Introduction

The MODerate-resolution Imaging Spectrometer (MODIS) is a passive Earth-radiation sensor scheduled for launch on the Earth Observing System (EOS) orbiting platform in 1998. MODIS senses reflected solar radiation during daylight hours and Earth-emitted thermal radiation (infrared) continuously (day and night).

Science products for the MODIS instrument will be developed and validated by a team of twenty-four Earth scientists selected for their expertise in instrument calibration, atmospheric science, ocean science, and land science. Since the team members were chosen for their scientific expertise, the team includes members with varying interest and abilities in data system implementation. To accommodate the individual differences, the MODIS Team Leader is allowing the science team members to themselves specify the extent to which they will develop the software they deliver to the project. Some team members may want to simply specify the equations to be used for product generation without delivering any processing code. Other team members may deliver prototype code that runs on the scientists home computing facility, and still others may deliver full-up code, ready for operational use on the designated high-speed processing facilities.

Some required MODIS processing tasks are not included in any Science Team Member's domain of interest (e.g., basic MODIS Level-1 instrument data processing). To develop code to do tasks not supported by anyone on the science team and to assist in porting scientist's code to operational data production facilities, the MODIS Team Leader has designated a software support group called the MODIS Science Data Support Team (SDST). This document describes the functions to be performed by the SDST, the required interfaces between the SDST and other MODIS and EOS data groups, and the specific hardware needed to support near-term SDST activities. This document presents an evolutionary approach to SDST facility development, and it contains a functional description of SDST facilities at each of several proposed phases of evolution.

2.0 The Team Leader SCF Environment

The overall data system that supports the EOS program is called the EOS Data and Information System (EOSDIS). The EOSDIS includes data communications components that handle data transfer to and from the platform as well as other components that generate the commands to be transferred to the platform and interpret the data received from the platform. Instrument command generation and the scientific interpretation, storage, and distribution of EOS data will be done in a subset of the EOSDIS called the EOSDIS Core System (ECS). See Figure 1. The ECS provides an Instrument Command Center (ICC) for each individual instrument. Operational processing of instrument data to generate Earth-science products will be done in a sub-facility of the ECS called the Product Generation System (PGS), and storage and distribution of data will be done in another facility called the Data Archive and Distribution System (DADS). The data user interface to the DADS is handled by the Information Management System (IMS). The basic structure is as indicated in Figure 1. The ECS also contains other components that do not interface with MODIS processing.

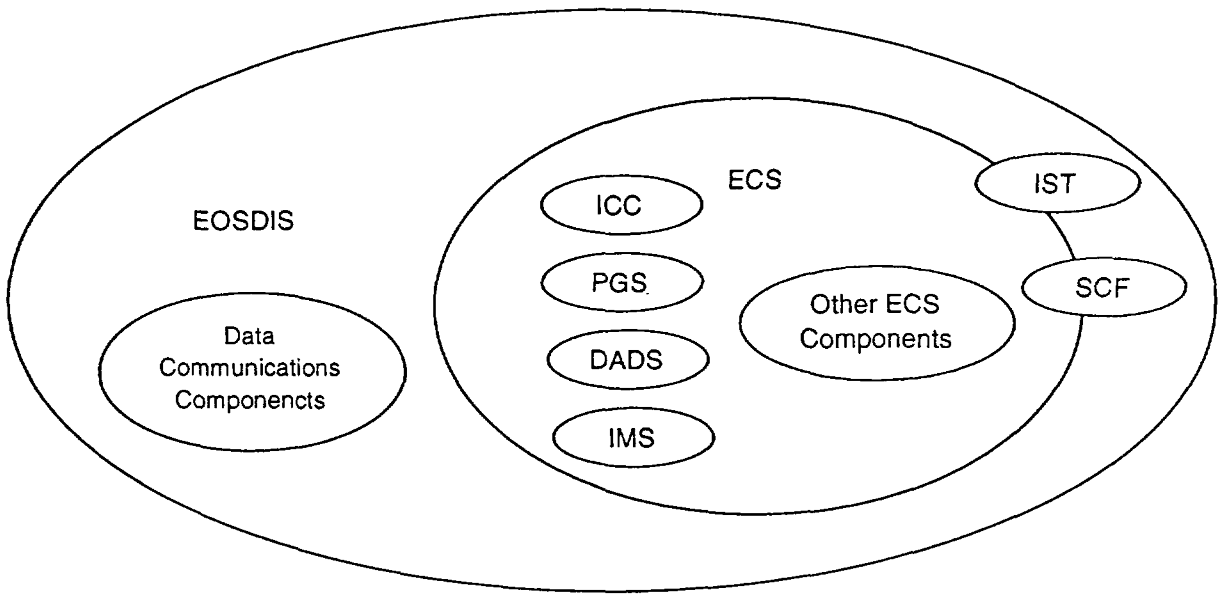


Figure 1. EOSDIS Structure and ECS Components Interfacing with MODIS Processing

To allow the Science Team Leader (and possibly other Science Team Members) to monitor instrument behavior and participate in instrument command decisions without being physically present at the ICC, the ECS will provide a software toolkit known as the Instrument Support Terminal (IST). The IST toolkit will run on a local terminal or workstation provided by the Team Leader, his designate, or other participating Team Members. The IST allows the Team Leader to interactively participate in instrument planning and scheduling, review engineering data, analyze instrument trends and investigate anomalies (as required), and interactively develop command requests.

The initial development of software to produce EOS Standard Products, the production of Special Data Products (products produced for a subset of the available data and not accepted for routine production on the PGS), the validation of Standard and Special Data Products, and research activities of the Science Team Members will be done independently at the individual scientist's home computing facility, called a Science Computing Facility (SCF). The relationship between ECS facilities and the ISTs and SCFs is defined in the ECS Specification, and this specification is the formal basis for many of the requirements and functional relationships cited in this document.

Besides basic IST and SCF functions related to instrument monitoring and control and the production and validation of science products, the MODIS Team Leader must also support other functions related to his unique position as team leader. To assist with these functions, the MODIS Team Leader has defined the three support groups shown in Figure 2. The SDST was discussed above. The MODIS Characterization Support Team (MCST) provides support related to monitoring and calibration of the MODIS instrument. The MCST is planning a near-real-time instrument monitoring effort that will examine segments of the MODIS instrument data as these data are returned from the observing platform. The MCST will do a number of instrument-related investigations and will use general purpose computing facilities as well as special purpose computers dedicated to the instrument monitoring task. The MODIS Administrative Support Team (MAST) will provide basic administrative support to the Team Leader and the Science Team and will use computers only for administrative tasks. In this document, the entire complement of computers and associated peripherals available to the Team Leader will be called the Team Leader Computing Facility (TLCF).

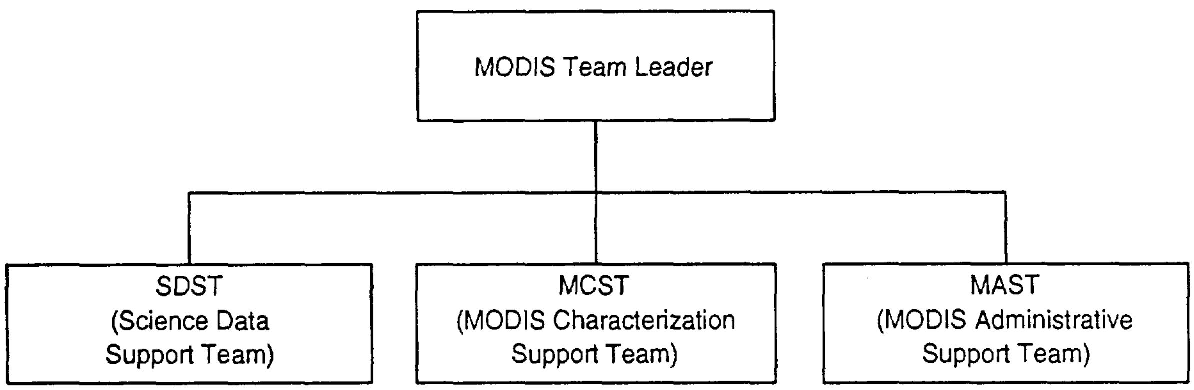


Figure 2. MODIS Support Teams.

The ECS specification defines the support that the TLCF must provide to the ECS. Information flows is both directions across the interface between the ECS and the Team Leader's facility, and the basic nature of the relationship is indicated in Figure 3. In this diagram, the TLCF has been broken into its two logical components: the Team Leader SCF, which supports software development and integration by the SDST and other Science Team Members, and the MCST SCF, which includes the MODIS IST, and which supports MCST activities. Although the SDST and MCST portions of the TLCF can be shown as logically separate entities, as in the diagram, it is expected that the SDST and MCST will share at least some physical facilities, i.e. at least some

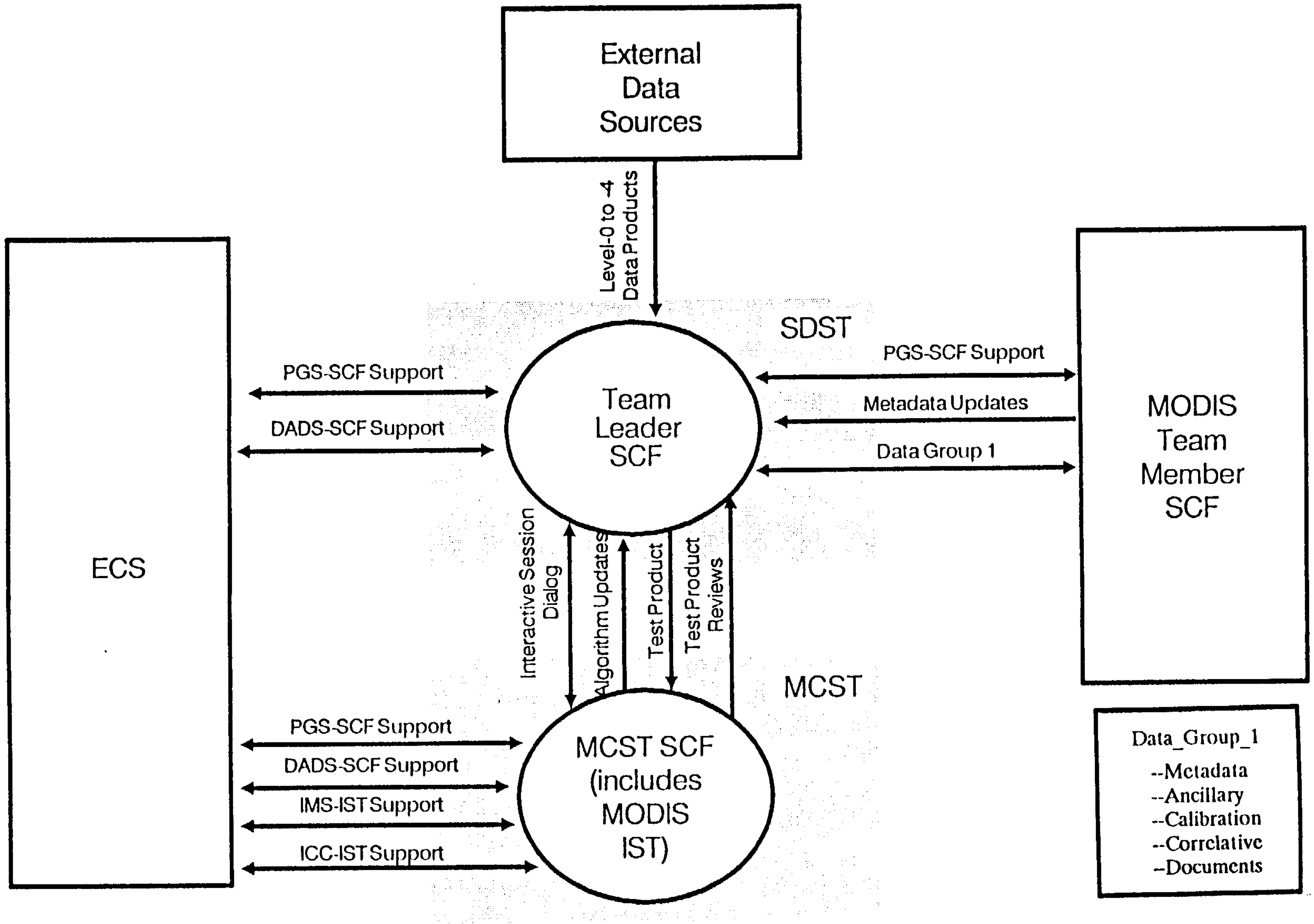


Figure 3. TLCF Data Flow Diagram Showing the Logical Components of the TLCF

portions of the Team Leader SCF and the MCST SCF will physically reside on the same computing facilities. To provide a high-level overview, the data flows between the ECS and the TLCF components (Team Leader SCF and MCST SCF) have been shown generically in the diagram, e.g. data flow between the ECS and the Team Leader SCF has been shown as "PGS-SCF Support" and "DADS-SCF Support". Diagrams showing the expanded definitions of data flows between the ECS and the TLCF are given in Figure 4. A data dictionary defining Team Leader SCF data flows is included as Appendix A to this document.

Figure 3 shows other data system entities that interface with the Team Leader SCF. "External Data Sources" will provide Level-0 through Level-4 data products for non-EOS instruments that are needed for MODIS algorithm development and product validation. This is a one-way data flow from the external source to the Team Leader SCF. It is the responsibility of the individual Science Team Leader or Member to initially obtain the external data required to develop and validate his product.

The interface shown with the "MODIS Team Member SCF" is more complicated. If the individual Team Member desires, the Team Leader SCF can serve as an intermediary between the Team Member and the PGS. Services to be provided by the Team Leader SCF include (potentially) code development for the individual Team Member, porting of data product code from the Team Member's SCF to PGS-compatible facilities, and the integration of multiple Team Member algorithms into a single, efficient, operational MODIS product generation system. Some MODIS products require other MODIS products as input and can only be generated after the required MODIS products for the corresponding area have been generated, and production scheduling is potentially a troublesome issue . Also, proper ordering of data product generation may minimize data input requirements and improve efficiency, i.e. if several algorithms requiring the same input data are run sequentially while the data is retained in memory, data is input only once for the entire procedure, and not once of each individual product algorithm.

If desired, the MODIS Team Member can interact directly with ECS facilities (specifically the PGS) to integrate and test his software; he need not utilize Team Leader services. This point is made clear in Figure 5, which shows the context diagram for the Team Member SCF and explicitly shows potential direct interaction between the SCF and the ECS as well as the potential utilization of Team Leader services.

3.0 SDST Facility Development Phases

The proposed SDST facility development schedule is primarily determined by one key requirement related to software integration and testing. Although the ECS is developing a PGS toolkit that is intended to simulate the operational features of the PGS at the scientist's local SCF, it is expected that full cross-platform code portability from the SCF to the PGS cannot be assured, and MODIS algorithm integration and testing at the Team Leader SCF will be done using facilities fully compatible with the operational PGS. Since algorithm integration and testing is critical to the

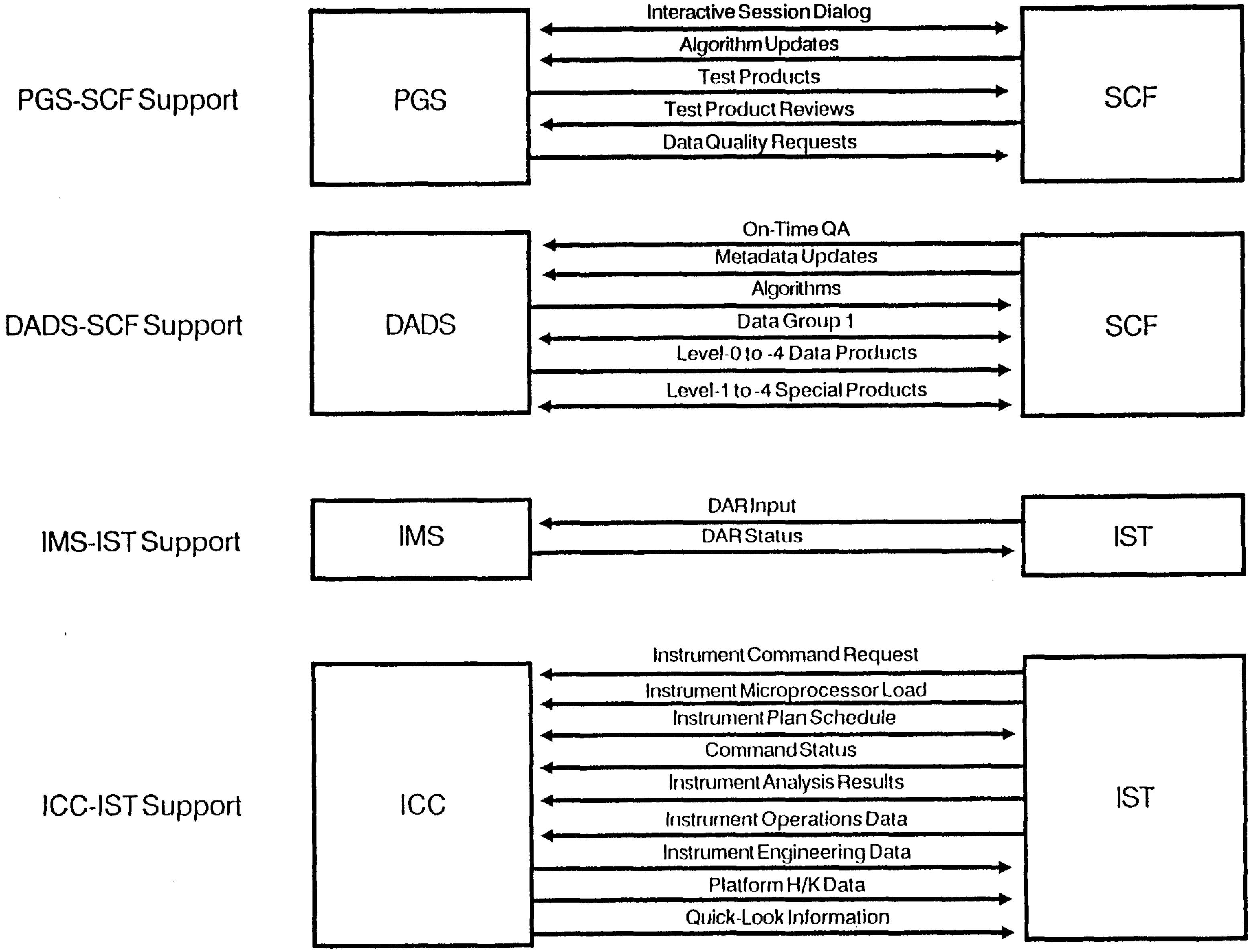


Figure 4. Expanded Definitions of Support Components

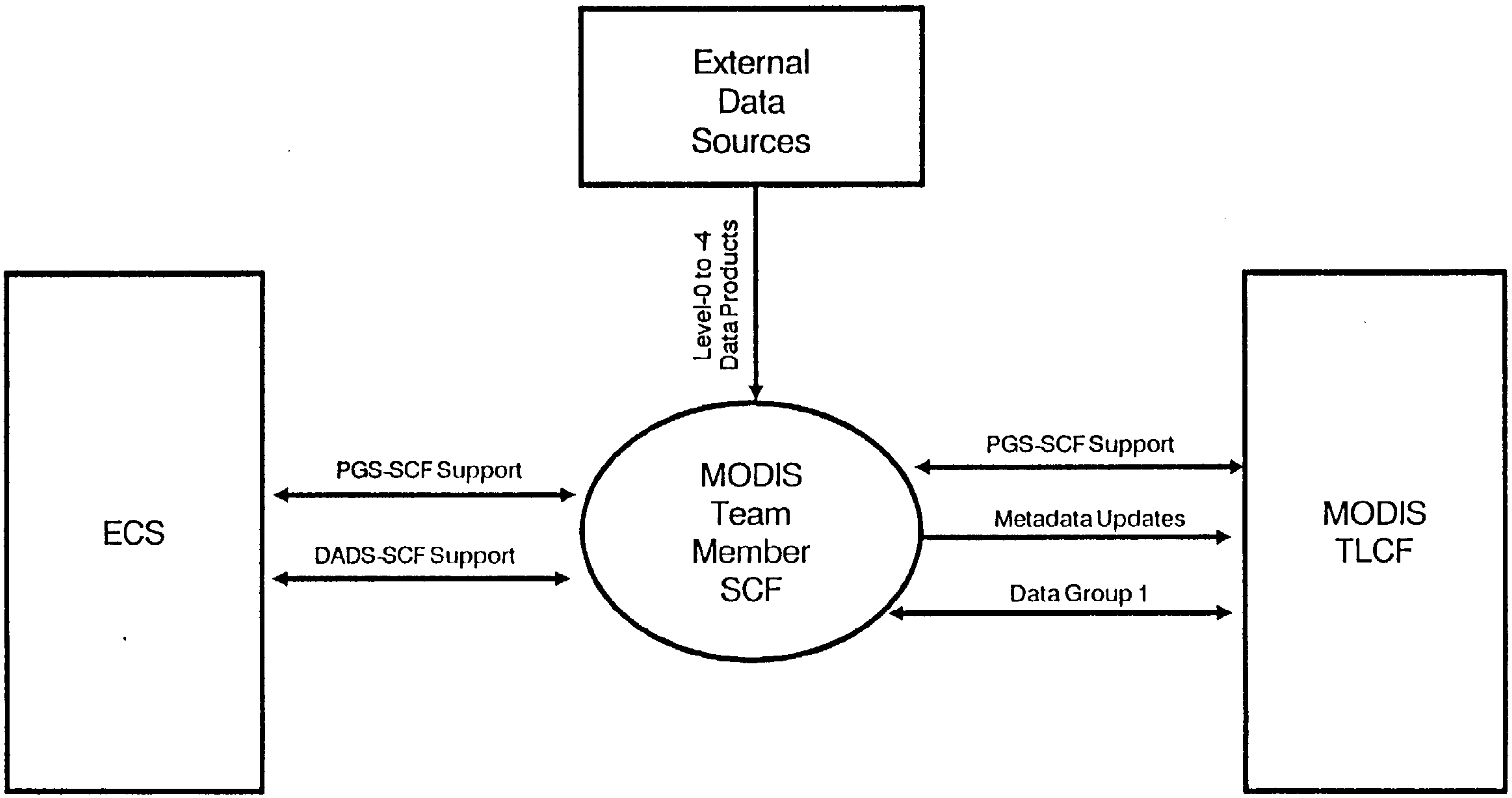


Figure 5. MODIS Team Member SCF Context Diagram

timely completion of MODIS processing software by the launch date, integration and testing must begin as soon as possible, and the critical event shaping the development schedule is the availability of PGS-compatible hardware and software for SDST use, along with algorithms from the Science Team (at least in prototype form). The ECS contract is not to be awarded until November, 1992; the PGS computing architecture may not be determined until perhaps a year later, and facility procurement will likely add at least another eight months of delay, so that, at best, a PGS-compatible facility may first be available in mid-1994. Software integration and testing is critical and should begin on that date or as soon thereafter as possible.

The proposed facility development schedule is shown in Figure 6. As stated, use of PGS-compatible or "mini-PGS" facilities (Phase II) should begin in mid-1994. Most Science Team Members will be only in the preliminary stages of testing at that time, and it is expected that a smaller PGS system, or "mini-PGS" will be adequate to meet requirements until perhaps a few years before launch, when development and testing efforts will become more intensive, and a full-up PGS-compatible system is required (Phase III). Present plans are to size the full-up Team Leader SCF to equal MODIS operational processing requirements at the PGS, i.e., in the absence of integration and test activity, the full-up facility should be capable of MODIS operational processing with no support from the PGS.

SDST Facility Development Phases							
	1992	1993	1994	1995	1996	1997	1998
Phase I—Unix Workstation							▾Launch
Phase II—Mini-PGS							
Phase III—Full-Up System							

Figure 6. SDST Facility Development Phases

Phase I is an interim phase that will allow the SDST to begin processing of prototype data sets obtained from MODIS precursor instruments, run Computer Aided Software Engineering (CASE) tools, develop MODIS Level-1 processing code, and begin integration testing on a preliminary set of prototype science algorithms. Many of the Science Team Members will be using Unix workstations, and for Phase I, the SDST will also use a Unix Workstation. Specific requirements for the workstation will be listed in Section 6.

4.0 Functional Requirements for SDST Facilities

Many of the required SDST functions can be inferred directly from the data flows shown in Figures 3 and 4.

"PGS-SCF Support" includes "Interactive Session Dialog" that supports general communication between the PGS and the SCF for software integration and test. "Algorithm Updates", "Test Products", and "Test Product Reviews" also support algorithm integration and test at the PGS. "Algorithm Updates" include the source code for the candidate algorithm, algorithm documentation, and a job step control skeleton that controls the execution sequence for the algorithm and the interchange of data with other programs being executed. Test products generated by the candidate algorithms are sent to the SCF. Reviews of the test products are sent back to the PGS. Algorithm development and maintenance is one of the primary functions performed at the SCFs.

"Data Quality Requests" originate at the PGS and are related to another SCF function, namely, validation of data products routinely produced at the PGS. A "Data Quality Request" includes a time window for Team Member response. "On-Time QA" [shown as part of the "DADS-SCF Support"] is a Team Member response received within the time limit. "On-Time QA" information is used to complete the QA fields of the product metadata as the product is shipped to the DADS. Responses received after the time window closes go directly to the DADS as "Metadata Updates", the next item shown in the "DADS-SCF Support".

"Algorithms", "Data Group 1" [see definition in Figure 3], "Level-0 to -4 Data Products", and "Level-1 to -4 Special Products" are related to another SCF function, namely, research investigations. If desired, the Team Member may access other scientist's algorithms stored at the DADS to support his own development efforts. Also to support his investigations at the local SCF, the Team Member may access "Data Group 1" items including Metadata on data items stored at the DADS, Ancillary, Calibration, and Correlative data, and algorithm documentation, as well as "Level-0 to -4 Data Products" for other instruments and "Level-1 to -4 Special Products" produced at other Team Member SCFs within the EOSDIS. Such investigations may or may not result in useful data products to be shared with other investigators. If not, the Team Member effort is a simple research investigation. If useful products are produced, these products are to be shared with other investigators and are known as Special Products. Special Products with their associated metadata and documentation are transferred from the SCF to the DADS in the "Data Group 1" and "Level-1 to -4 Special Product" flows going from the SCF to the DADS.

The "Level-0 to -4 Data Products" flow from "External Data Sources" recognizes the fact that not all data needed by a Team Member for a scientific investigation will be available from the DADS. The Team Member can best identify appropriate "External Data Sources" for his investigation.

The relationship shown between the "Team Leader SCF" and the "MODIS Team Member SCF" recognizes the potential support function that the "Team Leader SCF" may provide to other Science Team Members. Besides integration and testing support for MODIS algorithms, the Team Leader SCF may also perform routine QA of Team Member products, if the Team Member desires. These support functions are embodied in the "PGS-SCF Support" flow between the Team Leader SCF and the Team Member SCF. Also, if the Team Member desires, the Team Leader SCF may support the production of Special Products for the Team Member if, for example, the hardware capability of the scientist's local SCF is inadequate to support the desired volume of Special Product generation.

The Team Member will perform the QA of such products, and the Team Member will supply "Metadata Updates", as shown, to complete the QA field in the metadata for such Special Products. "Data Group 1" data flows are also related to potential Special Product generation at the Team Member SCF.

The relationship shown between the Team Leader SCF and the MCST SCF includes all aspects of the Team Leader relationship with any other SCF except that the Team Leader SCF is not likely to provide routine QA of data products for the MCST nor is it likely to produce Special Products for the MCST. The Team Leader SCF does support the integration and testing of MCST algorithms.

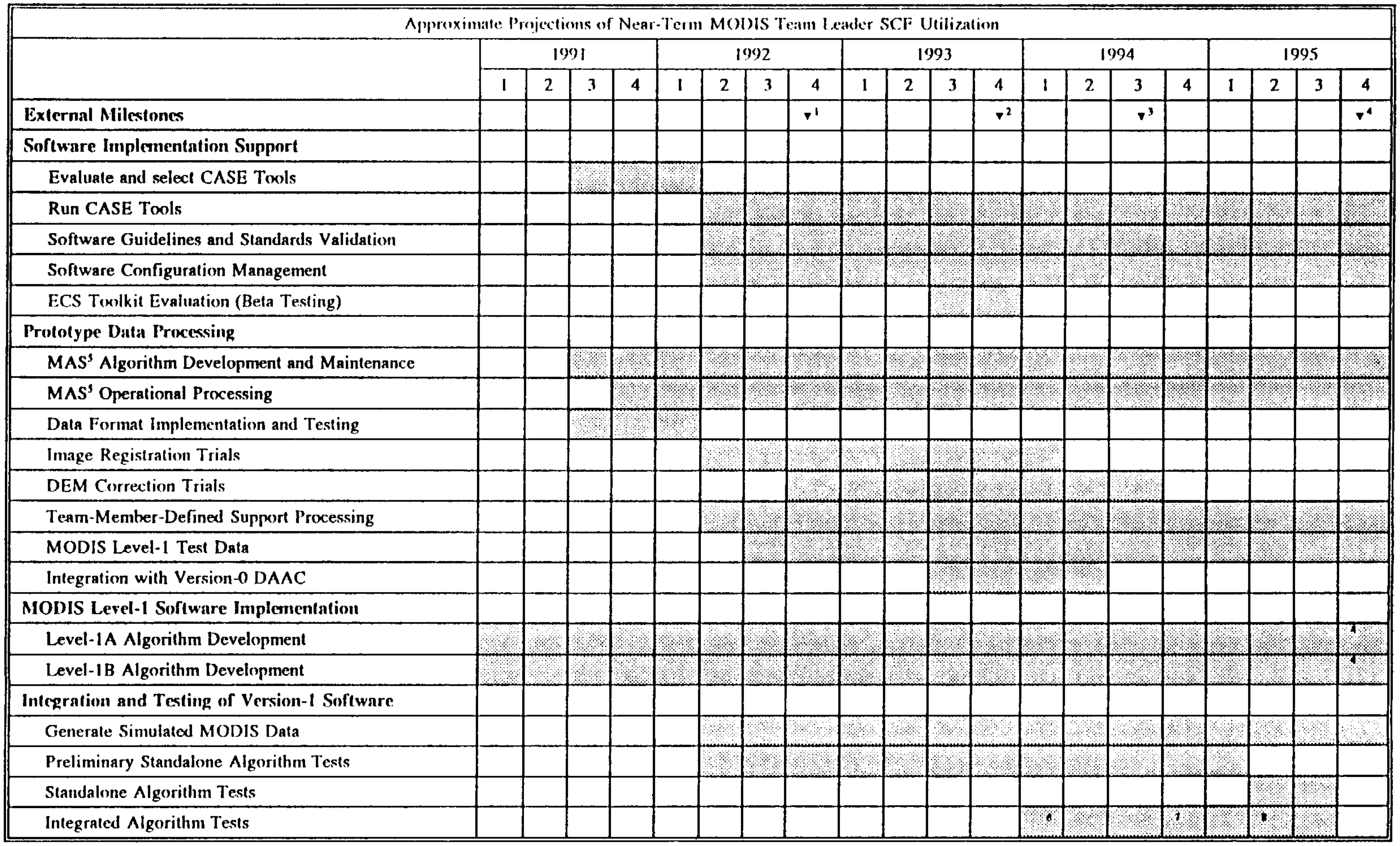
In addition to supporting formal functional relationships expressed in data flow diagrams and discussed above, the Team Leader SCF will also provide a number of short-term and special purpose support services for the MODIS Science Team. Figure 7 is a list of Team-Leader-unique support functions identified thus far in the effort. Since the Team Leader is responsible for providing services not otherwise provided within the Science Team, this list will doubtlessly evolve as implementation progresses and new needs are identified. For each function, the figure shows an associated time interval during which the support service is thought to be needed.

5.0 Networking and Communications Requirements for SDST Facilities

The basic networking and communications requirements for the Team Leader SCF are implicit in the data flow diagram shown in Figure 3. The Team Leader SCF interfaces with PGS and DADS components of the ECS, with "External Data Sources", with the respective SCFs of other MODIS Science Team Members, and with the MCST portions of the local TLCF. Of these, the Team Leader SCF, major portions of the MODIS PGS and DADS, the MCST SCF, and several MODIS Team Member SCFs are expected to be located at Goddard Space Flight Center (GSFC) and can use presently-existing (during Phase I) and enhanced (for Phase II and III) networking capability provided for EOS use at the Center.

The ECS-Team Leader SCF Interface.

All MODIS Level-1 Data Products will be produced and stored at GSFC. The production and storage of MODIS Level-2 through Level-4 products will be distributed across three data centers as shown in Table 1. Besides GSFC, the contributing centers are the Earth Resources Observation System (EROS) Data Center (EDC) in Sioux Falls, South Dakota and the National Snow and Ice Data Center (NSIDC) in Boulder, Colorado. All MODIS atmospheric and ocean products will be produced and stored at GSFC. Level-2 land products will also be produced at GSFC. Level-3 and 4 land data products will be produced at EDC and Level-2 through Level-4 snow and ice products will be produced at NSIDC.



[1]ECS Contract Award
[2]PDR-PGS Architecture Chosen (approx.)
[3]PGS-compatible machine delivered (approx.)
[4]Version 1 software due
[5]MODIS Airborne Simulator (MAS) or a successor instrument
[6]Team Member Version 0 code at SDST for integration and testing (with selected TMs, not contractually required), simulated data needed
[7]Review progress, make changes
[8]Team Member Version 1 code due at SDST for integration and testing

Figure 7. Approximate Projections of Near-Term MODIS Team Leader SCF Utilization

Table 1
Production and Storage of MODIS Level-2 through Level-4* Data Products

		Level-2	Level-3	Level-4
Atmospheric	PGS DADS	GSFC GSFC	GSFC GSFC	GSFC GSFC
Ocean	PGS DADS	GSFC GSFC	GSFC GSFC	GSFC GSFC
Land	PGS DADS	GSFC EDC	EDC EDC	EDC EDC
Snow/Ice	PGS DADS	NSIDC NSIDC	NSIDC NSIDC	NSIDC NSIDC
*All MODIS Level-1 Data Products are produced and stored at GSFC.				

Figure 8 shows a breakout of the ECS-Team Leader SCF interface to show the separate PGS and DADS components at the individual data centers. The DADS is a distributed system and all DADS functions are accessible at any of the centers so that the Team Leader SCF at Goddard can access all DADS-supported functions for all the data centers at the local GSFC DADS. This relationship among the individual DADS components is indicated by the connecting arrows shown on the left in the figure; special MODIS communications are not required to support this DADS function.

Besides the local links within GSFC, two distant link requirements remain for the ECS-Team Leader SCF interface. PGS-SCF Support for Level-3 and 4 land products is required with the EDC in Sioux Falls, SD and PGS-SCF Support for Level-2, 3, and 4 snow and ice products is needed with the NSIDC in Boulder, CO. Examination of PGS-SCF Support as defined in Figure 4 reveals two basic functions that are involved: integration and testing of product algorithms at the remote sites and, potentially, routine QA of operational products produced at the remote sites (if the responsible Team Member requests Team Leader assistance with this task). Communications to support integration and testing will be needed only sporadically and will likely involve only relatively small volumes of data to be transferred. Although the communications requirement for routine QA is potentially larger, it is thought that most Science Team Members will not want to examine large volumes of their products at the SCF, and therefore, data volume for this function will also be moderate.

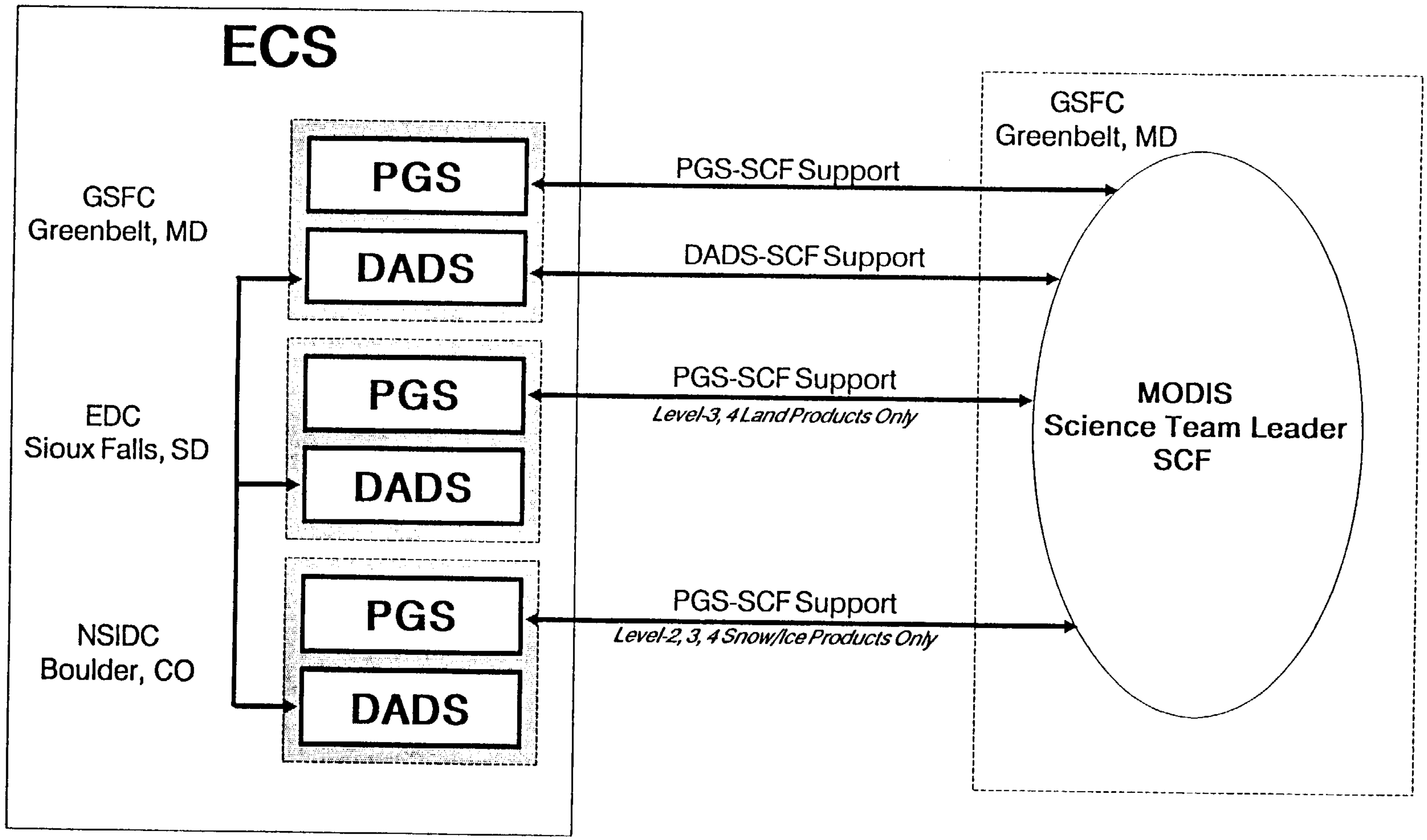


Figure 8. Expanded View of ECS-MODIS Team Leader SCF Interface

The "External Data Source"-Team Leader SCF Interface

SCF communications with "External Data Sources" could potentially involve a large segment of the worldwide Earth-science community. The EOSDIS Science Network (ESN) (to be developed by the ECS contractor) will provide gateway access to the NASA Science Internet (NSI), which will, in turn, provide the required access to the worldwide community. The MODIS project will not require special communications support once the services of ESN become available. In the near-term, the SCF will also have access to internet services, and it is expected that most required access to "External Data Sources" can be handled via existing internet services.

The Team Leader SCF-MODIS Team Member SCF Interface.

Because of the unique roles dicsussed above, the Team Leader SCF is likely to have the largest intra-team communication requirement of any MODIS Team Member. Although data transfer volumes for the Team Leader SCF could be appreciable, most of the functions supported are not operationally pressing, and short delays in communications response may be tolerable. Since most of the ocean product code is being developed and integrated at the University of Miami, communications requirements with that facility may be particularly large. Intra-team communications requirements should be reanalyzed as Phase II facilities are acquired and ESN communication services to the SCFs are implemented. In the near-term, most Team Leader communications with MODIS Team Members can be handled via internet or other presently-existing data networks. Six of the MODIS Team Members have *not* indicated that they have access to internet and a few high-speed modem connections are planned to provide backup access. [Specifics of Phase I communications requirements are listed below.]

The Team Leader SCF-MCST SCF Interface.

The Team Leader SCF and the MCST SCF are both components of the TLCF and will likely share at least some physical facilities at GSFC. For the near-term (Phase I), it appears that data communications requirements between these components can be adequately handled by the existing GSFC network. In the long-term (Phase II and III), alternative communications will doubtlessly be implemented as part of the ESN. Perhaps a fiber optic link (FDDI) will be appropriate.

Near-Term Team Leader SCF Communications Requirements.

The list of near-term SCF functions given in Figure 7 has been examined to extract those functions requiring communications support. The resulting list of functions and communications requirements is given in Table 2. The near-term functions requiring communications support include the remote use of CASE tools, "Beta" testing of ECS-provided toolkits, Team-Member-defined support / integration and testing support, and integration of MODIS Airborne Simulator (MAS) and other prototype MODIS data sets with the Version 0 EOSDIS processing system. Proposed communications support is listed for each requirement. Figures 8-13 show the software hierarchy required to support the required communications. Besides Unix, TCP/IP and X-Window

Table 2
Near-Term Phase I Communications Requirements for the MODIS Team Leader SCF

Function	User	Remote Site	Environment	Protocol	Medium	Rate (Kbps)
Run CASE Tools	SDST	Terminal Room	X-Windows/MOTIF	TCP/IP	Goddard Network	
	TMs	SCFs	X-Windows/MOTIF	TCP/IP	Internet	
	TMs	SCFs	X-Windows/MOTIF	SLIP/CSLIP	Phone Line (1)	14.4, V32 bis, V42 bis
ECS Toolkit Evaluation (Beta Testing)						
ESN Toolkit	SDST	PGS			ESN	
SMC (CASE) Toolkit	SDST	PGS			ESN	
IMS Toolkit	SDST	PGS			ESN	
	SDST	Anywhere		TCP/IP	Internet	
	SDST	Anywhere			Phone Line (1)	14.4, V32 bis, V42 bis
Team-Member-Defined Support/ Integration and Testing Support	TMs	SCFs	TELNET X-Windows/MOTIF	TCP/IP	Internet	
	TMs	SCFs	TELNET X-Windows/MOTIF	SLIP/CSLIP	Phone Line (1)	14.4 V32 bis, V42 bis
Integration with Version-0	SDST	DAAC	TELNET X-Windows/MOTIF			

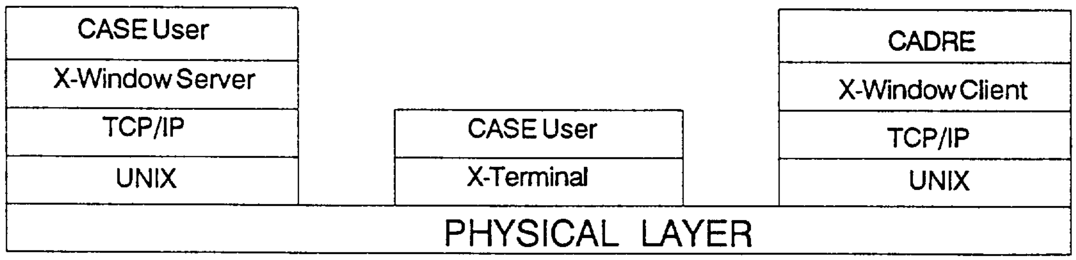


Figure 8. CASE Tool Operation

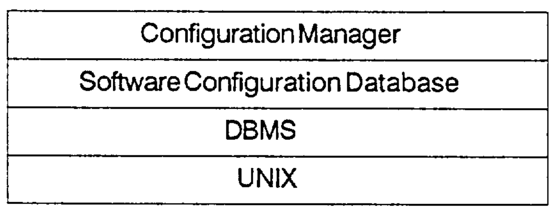


Figure 9. Software Configuration Management

Figure 10. ECS Toolkit Evaluation

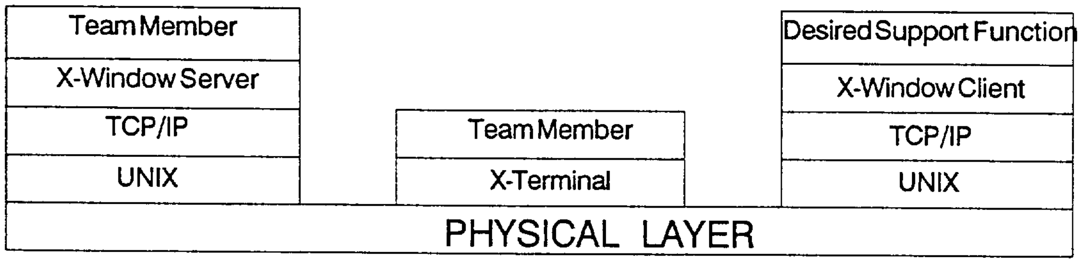


Figure 11. Team Member Support (image-based)

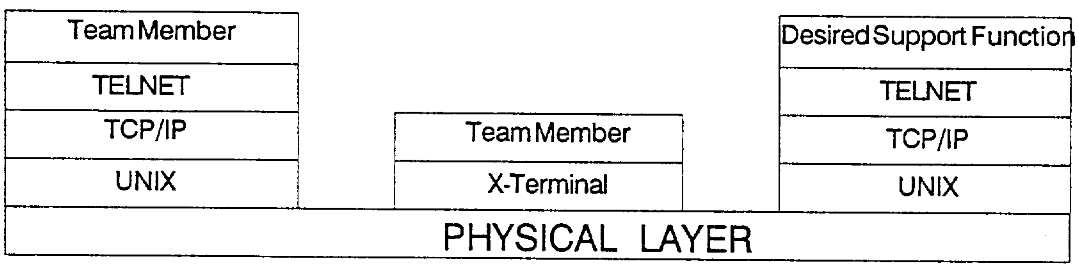


Figure 12. Team Member Support (character-based)

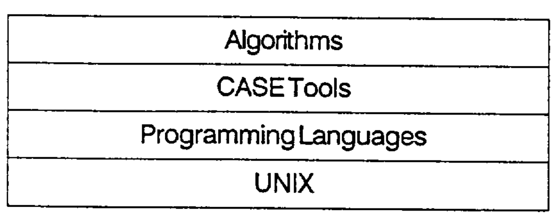


Figure 13. Algorithm Development and Product Generation

support (at least X-11, Revision 4) is essential.

6.0 The Proposed Phase I (Near-Term) System

Specific plans have been made to procure a Phase I (near-term) system for Team Leader SCF use. The system is to include:

- A Unix Workstation
- 178 MB of Random Access Memory
- 2.3 GB of Hard Disk Storage
- A CD-ROM Reader
- Black/White Postscript Printer*
- Color Postscript Printer*
- 9-Track 6250 bpi Magnetic Tape Drive (2)*
- 8 mm Tape Drive*

Items marked with a (*) are already available. A performance specification for the Unix workstation has been developed; the workstation shall:

- Support ten X-Windows users with no degradation in performance
- Perform at 70 SPECmarks and 20 MFLOPS (double-precision Linpack)
- Offer X-Windows performance of at least one million vectors/sec and 250,000 characters /sec
- Support single and differential SCSI II devices with I/O burst transfer rates of 5 MB/sec and 10 MB/sec, respectively, with sustained transfer rates of at least 2 MB/sec
- Run the Hewlett-Packard *Softbench* development environment (Sun, HP, and IBM workstations run this software), Cadre's *Teamwork* structured design tools (Sun, HP, IBM, and DEC workstations and IBM PCs run this software) and *QA C*
- Run the OSF Motif graphical user interface and the KORN shell
- Support a serial protocol, SLIP and/or PPP

The software components HP *Softbench*, Cadre *Teamwork*, and *QA C* are to be procured with the initial system. Cost figures for the proposed hardware and software configuration have been developed and are included as Appendix B in this report.

Appendix A

Data Dictionary for Team Leader SCF Data Flows

ALGORITHMS consist of the executable programs for science product generation, source code of these executable programs, job control scripts, and algorithm documentation. Algorithms are the result of new or updated science algorithms passing through the integration and test process, involving the scientist and the PGS's algorithm integration and test staff. After formal approval, algorithms are delivered by the PGS to the DADS for storage, and are retrieved as needed to support product production. The DADS shall also archive algorithms contributed as EOSDIS resources by other data centers. Algorithms shall be orderable and distributed to authorized users. Some frequently used algorithms may also be kept on line in the PGS.

ALGORITHM UPDATES are delivered to the PGS's integration and test environment by scientists at an SCF. They represent changes to existing production algorithms, or a new algorithm to produce a new Standard Product. Algorithm updates include the source code for the candidate algorithm, its associated documentation, and a job step control skeleton. The source code will be compiled to form an executable program suite as part of the integration and test process. The job step control skeleton contains instructions that control the sequence of execution of, and the interchange of data between programs from the executable program suite. Test data sets and calibration data should also be included.

ANCILLARY DATA refers to any data, other than Standard Products, that are required as input in the generation of a Standard Product. This may include selected engineering data from the EOS platform, ephemeris data, as well as non-EOS ancillary data. All ancillary data is received by the PGS from the DADS.

CALIBRATION is the collection of data required to perform calibration of the instrument science data, instrument engineering data, and the spacecraft or platform engineering data. It includes pre-flight calibration measurements, in-flight calibrator measurements, calibration equation coefficients derived from calibration software routines, and ground truth data that are to be used in the data calibration processing routine.

CORRELATIVE data are scientific data needed to evaluate and validate EOS data products.

DATA QUALITY REQUEST is a request issued by the PGS to a scientist at an SCF to perform QA of a particular product before future processing or distribution. A time window is applied to the request in keeping with the production schedule.

DOCUMENTS are the hardcopy or digitized references or records about an instrument or the products generated from its data. These shall be archived at the DADS.

INTERACTIVE SESSION DIALOG consists of messages that flow between a scientist at an SCF and the PGS that support general communication with the Integration and Test Service. This includes logins, mail messages, etc.

L0-L4 DATA PRODUCTS consist of L0 Data Products from the IPs, the ADCs and ODCs, and L1-L4 Standard Products produced in the PGS.

L1-L4 SPECIAL PRODUCTS are special science data products consisting of L1A, L1B, L2, L3, and L4 which are produced at the SCFs. These shall be archived at the DADS and distributed to authorized requestors.

METADATA is data which describes the content, format, and utility of a Standard Product. It includes standard metadata (i.e., algorithm and calibration numbers, size of product, date created, etc.), algorithm-derived metadata, QA information from the PI's, summary statistics and an audit trail. Metadata is received by each DADS with the corresponding data sets. DADS validates it physically, updates it with inventory information, enters it into a distributed database (to which the IMS has access), and archives it. Metadata about special products produced at SCF shall be sent to DADS along with their associated data products.

METADATA UPDATES are additional or changed metadata items relating to a previously delivered product.

ON TIME QA is a response to a data quality request that is received within the established production time window. It is received from a scientist at an SCF. It consists of data which will be used to complete the QA fields of the metadata. Overdue QA responses are sent directly to the DADS.

TEST PRODUCTS are science products generated by new or updated algorithms during the integration and test period. Test products are delivered to scientists at an SCF.

TEST PRODUCT REVIEWS are evaluations of test products that are used to determine how to proceed in the integration and test process for a new or updated algorithm. A review may indicate the need for further algorithm refinement, or it may indicate that a candidate algorithm is ready for formal adoption into the production environment. Test product reviews are received by the PGS from scientists at an SCF.

Appendix B

Facility Costs for FY 1992